

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-290788

(43)Date of publication of application : 19.10.2001

(51)Int.Cl.

G06F 15/177

G06F 9/46

G06F 11/34

G06F 13/00

(21)Application number : 2000-108443

(71)Applicant : MITSUBISHI ELECTRIC CORP

(22)Date of filing : 10.04.2000

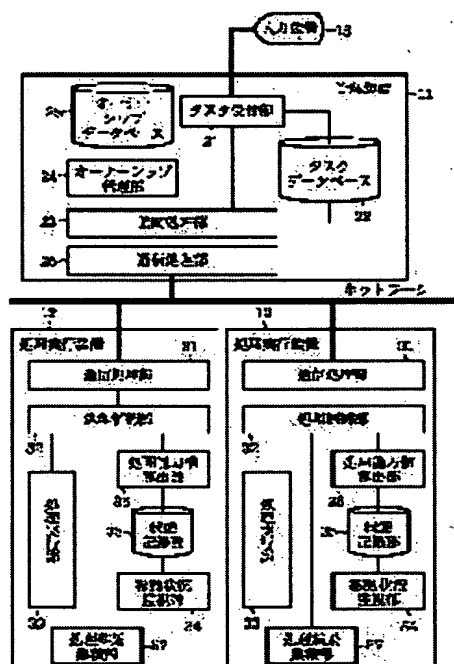
(72)Inventor : TOKUMOTO SHUICHI

(54) DISTRIBUTED PROCESSING SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To suppress the increase of a communication rate even if the number of task processings and the number of execution computers are increased.

SOLUTION: The throughput value calculating part 36 of a processing execution device 12 refers to operation data stored in a state recording part 35 and calculates the latest throughput value associated with a task during a processing at a prescribed interval. The processing execution device 12 transmits the throughput value to the cooperation processing part 23 of a managing device 11. When a task reception part 21 receives the processing instruction of the task, the cooperation processing part 23 decides the processing execution device 12 executing the processing of the received task based on the throughput value transmitted from the respective processing execution devices 12.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-290788

(P2001-290788A)

(43) 公開日 平成13年10月19日 (2001. 10. 19)

(51) IntCl. ⁷	識別記号	F I	テーマコード* (参考)
G 0 6 F 15/177	6 7 4	G 0 6 F 15/177	6 7 4 B 5 B 0 4 2
9/46	3 6 0	9/46	3 6 0 C 5 B 0 4 5
11/34		11/34	S 5 B 0 8 9
13/00	3 5 7	13/00	3 5 7 Z 5 B 0 9 8

審査請求 未請求 請求項の数10 O L (全 14 頁)

(21) 出願番号 特願2000-108443 (P2000-108443)

(22) 出願日 平成12年4月10日 (2000. 4. 10)

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 徳本 修一

東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内

(74) 代理人 100066474

弁理士 田澤 博昭 (外1名)

Fターム(参考) 5B042 GA12 JJ30 MC29

5B045 GC04 CG09 JJ02 JJ08

5B089 GA11 JA11 KA06 KA16 KC30

MA03 MA07

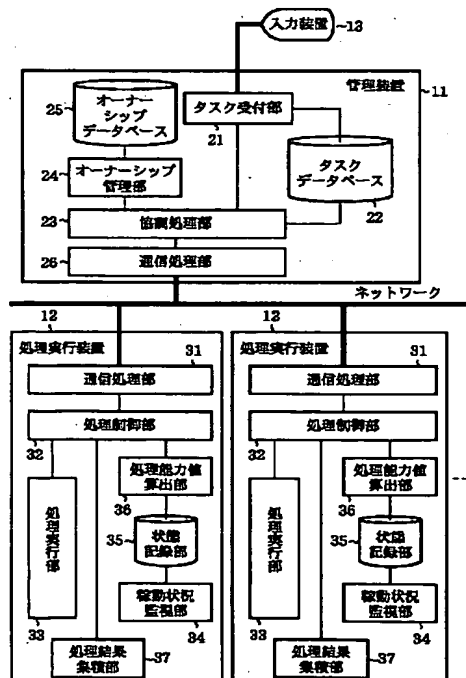
5B098 AA10 GA04 GC10 GD02

(54) 【発明の名称】 分散処理システム

(57) 【要約】

【課題】 タスク処理数や実行計算機数が増加しても通信料の増大を抑制する。

【解決手段】 処理実行装置12の処理能力値算出部36は、状態記録部35に記憶されている稼動データを参照し、所定の間隔で処理中のタスクと関連付けた最新の処理能力値を算出する。処理実行装置12は管理装置11の協調処理部23に処理能力値を送信する。タスク受付部21がタスクの処理命令を受け付けると、協調処理部23は、各処理実行装置12から送信された処理能力値に基づき、受け付けたタスクの処理を実行する処理実行装置12を決定する。



【 特許請求の範囲】

【 請求項1 】 タスクの処理を依頼する管理装置と、上記管理装置とネットワークにより接続され、上記管理装置からの依頼に基づきタスクの処理を実行する複数の処理実行装置とを備えた分散処理システムにおいて、

上記管理装置は、

外部よりタスクの処理命令を受け付けるタスク受付部と、

逐次把握している上記各処理実行装置のタスクと関連付けた処理状況を示す処理能力値に基づき、上記タスク受付部が受け付けたタスクの処理を実行する上記処理実行装置を決定して、タスクの実行管理権であるオーナーシップを設定し、決定した上記処理実行装置に、受け付けたタスクの処理依頼を行う 協調処理部とを備え、

上記処理実行装置は、

上記管理装置から依頼されたタスクの処理を実行する処理実行部と、

上記処理実行部によるタスクの処理における稼動データを計測する稼動状況監視部と、

上記稼動状況監視部が計測した稼動データを記録する状態記録部と、

上記状態記録部に記憶されている稼動データを参照し、所定の間隔で処理中のタスクと関連付けた最新の処理状況を示す処理能力値を算出する処理能力値算出部と、上記処理能力値算出部が算出した最新の処理能力値を上記管理装置に送信する処理制御部とを備えたことを特徴とする分散処理システム。

【 請求項2 】 協調処理部は、

各処理実行装置に対応して備えられ、対応する処理実行装置の処理能力値に基づき、受け付けたタスクの処理の実行可否を判断するノードと、

上記各ノードから報告された、タスクの処理の実行可否の判断結果と、処理能力値に基づき、オーナーシップの譲渡先を決定し、上記各ノードに通知するマネージャとを備えたことを特徴とする請求項1記載の分散処理システム。

【 請求項3 】 ノードは、

処理実行装置から送信された処理能力値を格納する処理能力データベースと、

上記処理能力データベースに格納されている処理能力値に基づき、受け付けたタスクの処理の実行可否を判断する状況分析部とを備えたことを特徴とする請求項2記載の分散処理システム。

【 請求項4 】 状況分析部は、各処理実行装置に固有な処理実行可能閾値を設定し、処理能力データベースに格納されている処理能力値に基づき、現在の計算機資源の占有率を求め、受け付けたタスクを実行したときの予測占有率を計算し、上記計算機資源の占有率と上記予測占有率の合計値と、上記処理実行可能閾値とを比較することにより、受け付けたタスクの処理の実行可否を判断す

ることを特徴とする請求項3記載の分散処理システム。

【 請求項5 】 オーナーシップを譲渡されなかったノードが、オーナーシップを譲渡されたノードからの処理能力値に基づき、オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理状況を観測し、オーナーシップを譲渡されなかった処理実行装置が代わりにタスクの処理を実行した方が処理効率が良いと判断した場合に、マネージャにオーナーシップの変更要求を行うことを特徴とする請求項2記載の分散処理システム。

【 請求項6 】 マネージャがオーナーシップの変更指示を各ノードに行った場合に、タスクの処理を実行していた処理実行装置は、処理済の中間データを、オーナーシップを新しく譲渡された処理実行装置に送信し、上記オーナーシップを新しく譲渡された処理実行装置が、送信された中間データに基づき、タスクの処理を継続することを特徴とする請求項5記載の分散処理システム。

【 請求項7 】 オーナーシップを譲渡されなかったノードが、

オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理能力値により実行モデルを作成し、作成した実行モデルのモデル情報を更新し、自己の処理能力値と比較して、処理効率の評価を行うモデリング部を備えたことを特徴とする請求項5記載の分散処理システム。

【 請求項8 】 オーナーシップを譲渡されなかったノードが、

モデリング部が作成した実行モデルのモデル情報の変化率、変化則を算出し、所定の時間内の処理能力値を予測すると共に、所定の時間内の自己の処理能力値の予測を行う予測部を備え、

モデリング部が、上記予測部が予測した処理能力値を比較することにより、処理効率の評価を行うことを特徴とする請求項7記載の分散処理システム。

【 請求項9 】 ノードが、

システム全体の処理実行装置の処理能力値により、各実行モデルを作成し、作成した各実行モデルのモデル情報を更新するモデリング部と、

上記モデリング部が作成した実行モデルのモデル情報を更新する度に、全実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、

上記負荷判定部の判定結果に基づき、負荷を調整する代替案を作成する負荷調整部とを備え、

上記負荷判定部が、上記負荷調整部により作成された代替案を評価し、マネージャにオーナーシップの変更を要求することを特徴とする請求項2記載の分散処理システム。

【 請求項10 】 ノードが、オーナーシップを譲渡されなかった処理実行装置の処理能力値により、他の実行モデルを作成すると共に、自己の処理能力値から自己の実行モデルを作成し、上記他の実行モデルのモデル情報

と、上記自己の実行モデルのモデル情報を更新するモデリング部と、

上記モデリング部が作成した各実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、

上記負荷判定部の判定結果に基づき、負荷を調整するために処理能力の高い処理実行装置に対応するノードにタスク処理の依頼を行う外部依頼部とを備えたことを特徴とする請求項2記載の分散処理システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、ネットワーク上に構成される複数の計算機のうち、任意の計算機に処理が集中する不具合を回避する分散処理システムに関するものである。

【0002】

【従来の技術】図15は、例えば特開平10-207847号公報に示された従来の分散処理システムの構成を示すブロック図であり、図において、100はタスクの処理を実行させる実行用計算機200を決定する管理用計算機、200は、ネットワーク上に複数構成され、管理用計算機100の指示に基づき、タスクの処理を実行する実行用計算機である。

【0003】また、図15の管理用計算機100において、101はジョブ入力を受け付けるジョブ受付部、102は入力されたジョブに関する情報であるジョブデータを実行用計算機200に配信するジョブデータ配信部、103は、実行用計算機200からの情報をもとに、そのジョブをどの実行用計算機200に割り当てるかを決定し、割り当て作業を行う判定/分配部、104は実行用計算機200との通信処理を行う通信処理部である。111は管理する実行用計算機200のリストで、ディスク等に記憶されている。

【0004】さらに、図15の実行用計算機200において、201は管理用計算機100との通信処理を行う通信処理部、202は自身の稼動状況を定期的に収集してデータベース203に蓄積する稼動状況監視部、204は、管理用計算機100から受信したジョブ情報と、データベース203に蓄積されている稼動状況に関する情報から、通知されたジョブの現時点での実行能力を評価する実行能力値算出部、205は割り当てられたジョブを実行する実行部である。

【0005】次に動作について説明する。まず、管理用計算機100のジョブ受付部101がタスクの処理命令を受けると、ジョブデータ配信部102は、そのタスク情報(タスクの処理を実行するために必要な機能)を、通信処理部104を介してネットワークに接続されている全ての実行用計算機200に送信する。各実行用計算機200において、通信処理部201が管理用計算機100からタスク情報を受けると、実行能力値算出部204は、そのタスク情報と、データベース203に記憶さ

れている現在の処理状況を考慮して、そのタスクを実行することが可能であるか否かを評価し、その評価結果を、通信処理部201を介して管理用計算機100に送信する。

【0006】管理用計算機100は、任意の実行用計算機200に処理が集中する不具合を回避するために、ネットワークに接続されている全ての実行用計算機200から評価結果を定期的に受けると、判定/分配部103は、それらの評価結果を内部ルールに適用して、当該タスクの処理を実行させる実行用計算機200を決定する。

【0007】

【発明が解決しようとする課題】従来の分散処理システムは以上のように構成されているので、管理用計算機100がタスクの処理命令を受け付ける毎に、ネットワークに接続されている全ての実行用計算機200に対してタスク情報を送信して、全ての実行用計算機200から評価結果を受信しなければ、タスクを実行させる実行用計算機200を決定することができず、タスクの処理数や実行用計算機200の接続数が増加した場合には、通信料が増大するという課題があった。

【0008】また、従来の分散処理システムでのタスクの割り当て方法では、一度、実行用計算機200に割り当てたタスクは、タスクの処理中に処理先の実行用計算機200を変更することができず、実行用計算機200の動的な負荷の変化に対応することができないという課題があった。

【0009】この発明は上記のような課題を解決するためになされたもので、タスクの処理数や実行用計算機200の接続数が増加しても、通信料の増大を抑制することができると共に、分散処理システム内の実行用計算機200の負荷変化に対応し、システム全体における処理の効率向上を可能にする分散処理システムを得ることを目的とする。

【0010】

【課題を解決するための手段】この発明に係る分散処理システムは、タスクの処理を依頼する管理装置と、上記管理装置とネットワークにより接続され、上記管理装置からの依頼に基づきタスクの処理を実行する複数の処理実行装置とを備えたものにおいて、上記管理装置は、外部よりタスクの処理命令を受け付けるタスク受付部と、逐次把握している上記各処理実行装置のタスクと関連付けた処理状況を示す処理能力値に基づき、上記タスク受付部が受け付けたタスクの処理を実行する上記処理実行装置を決定して、タスクの実行管理権であるオーナーシップを設定し、決定した上記処理実行装置に、受け付けたタスクの処理依頼を行う協調処理部とを備え、上記処理実行装置は、上記管理装置から依頼されたタスクの処理を実行する処理実行部と、上記処理実行部によるタスクの処理における稼動データを計測する稼動状況監視部

10

20

30

40

50

と、上記稼働状況監視部が計測した稼働データを記録する状態記録部と、上記状態記録部に記憶されている稼働データを参照し、所定の間隔で処理中のタスクと関連付けた最新の処理状況を示す処理能力値を算出する処理能力値算出部と、上記処理能力値算出部が算出した最新の処理能力値を上記管理装置に送信する処理制御部とを備えたものである。

【0011】この発明に係る分散処理システムは、協調処理部が、各処理実行装置に対応して備えられ、対応する処理実行装置の処理能力値に基づき、受け付けたタスクの処理の実行可否を判断するノードと、上記各ノードから報告された、タスクの処理の実行可否の判断結果と、処理能力値に基づき、オーナーシップの譲渡先を決定し、上記各ノードに通知するマネージャとを備えたものである。

【0012】この発明に係る分散処理システムは、ノードが、処理実行装置から送信された処理能力値を格納する処理能力データベースと、上記処理能力データベースに格納されている処理能力値に基づき、受け付けたタスクの処理の実行可否を判断する状況分析部とを備えたものである。

【0013】この発明に係る分散処理システムは、状況分析部が、各処理実行装置に固有な処理実行可能閾値を設定し、処理能力データベースに格納されている処理能力値に基づき、現在の計算機資源の占有率を求め、受け付けたタスクを実行したときの予測占有率を計算し、上記計算機資源の占有率と上記予測占有率の合計値と、上記処理実行可能閾値とを比較することにより、受け付けたタスクの処理の実行可否を判断するものである。

【0014】この発明に係る分散処理システムは、オーナーシップを譲渡されなかったノードが、オーナーシップを譲渡されたノードからの処理能力値に基づき、オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理状況を観測し、オーナーシップを譲渡されなかった処理実行装置が代わりにタスクの処理を実行した方が処理効率が良いと判断した場合に、マネージャにオーナーシップの変更要求を行うものである。

【0015】この発明に係る分散処理システムは、マネージャがオーナーシップの変更指示を各ノードに行った場合に、タスクの処理を実行していた処理実行装置は、処理済の中間データを、オーナーシップを新しく譲渡された処理実行装置に送信し、上記オーナーシップを新しく譲渡された処理実行装置が、送信された中間データに基づき、タスクの処理を継続するものである。

【0016】この発明に係る分散処理システムは、オーナーシップを譲渡されなかったノードが、オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理能力値により実行モデルを作成し、作成した実行モデルのモデル情報を更新し、自己の処理能力値と比較して、処理効率の評価を行うモデリング部を備えたもの

である。

【0017】この発明に係る分散処理システムは、オーナーシップを譲渡されなかったノードが、モデリング部が作成した実行モデルのモデル情報の変化率、変化則を算出し、所定の時間内の処理能力値を予測すると共に、所定の時間内の自己の処理能力値の予測を行う予測部を備え、モデリング部が、上記予測部が予測した処理能力値を比較することにより、処理効率の評価を行うものである。

【0018】この発明に係る分散処理システムは、ノードが、システム全体の処理実行装置の処理能力値により、各実行モデルを作成し、作成した各実行モデルのモデル情報を更新するモデリング部と、上記モデリング部が作成した実行モデルのモデル情報を更新する度に、全実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、上記負荷判定部の判定結果に基づき、負荷を調整する代替案を作成する負荷調整部とを備え、上記負荷判定部が、上記負荷調整部により作成された代替案を評価し、マネージャにオーナーシップの変更を要求するものである。

【0019】この発明に係る分散処理システムは、ノードが、オーナーシップを譲渡されなかった処理実行装置の処理能力値により、他の実行モデルを作成すると共に、自己の処理能力値から自己の実行モデルを作成し、上記他の実行モデルのモデル情報と、上記自己の実行モデルのモデル情報を更新するモデリング部と、上記モデリング部が作成した各実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、上記負荷判定部の判定結果に基づき、負荷を調整するために処理能力の高い処理実行装置に対応するノードにタスク処理の依頼を行う外部依頼部とを備えたものである。

【0020】

【発明の実施の形態】以下、この発明の実施の一形態を説明する。

実施の形態1. 図1はこの発明の実施の形態1による分散処理システムの構成を示すブロック図であり、図において、11は、ネットワークに接続された複数の処理実行装置12のうち、タスクの処理を実行することが現在可能で、現在処理負荷が最小の処理実行装置12にタスクの処理を依頼する管理装置、12は管理装置11からタスクの処理依頼を受けると、そのタスクの処理を実行する処理実行装置、13は管理装置11にタスクの処理命令を入力する入力装置である。

【0021】また、図1の管理装置11において、21は外部の入力装置13と接続され、タスクの処理命令を受け付けるタスク受付部、22は受け付けたタスクの情報を格納するタスクデータベース、23はタスク受付部21がタスクの処理命令を受け付けたときに、各処理実行装置12の処理能力を参照して、現在、そのタスクの処理を実行する能力を有し、かつ、処理負担が最小の処

理実行装置12を検索する協調処理部である。

【0022】さらに、図1の管理装置11において、24はオーナーシップに関する情報を協調処理部23に提供すると共に、協調処理部23がオーナーシップの設定や変更を行うと、そのオーナーシップ情報を更新するオーナーシップ管理部、25はオーナーシップ情報を格納するオーナーシップデータベース、26は協調処理部23により検索された処理実行装置12に対してタスクの処理依頼を送信すると共に、各処理実行装置12から送信される処理能力値を受信して協調処理部に出力する通信処理部である。

【0023】図1の処理実行装置12において、31は、処理能力値算出部36で算出した処理能力値を管理装置11に送信すると共に、管理装置11から送信されるタスクの処理依頼を処理実行部33に出力する通信処理部、32は処理実行装置12の処理を制御する処理制御部、33は、管理装置11からタスクの処理依頼を受けると、そのタスクの処理を実行する処理実行部である。

【0024】また、図1の処理実行装置12において、34は処理実行部33によるタスクの処理の稼働状況を示す稼働データとして、現在のメモリ占有率と演算能力の消費率を定期的に計測する稼働状況監視部、35は稼働状況監視部34により計測された稼働データを格納する状態記録部、36は、状態記録部35に記録されている稼働データを参照し、処理中のタスク名、実行中の機能名と関連付けた最新の処理能力値を算出する処理能力値算出部である。37は処理実行部33による処理の中間結果を蓄積する処理結果集積部である。

【0025】図2は管理装置11における協調処理部23の内部構成を示すブロック図であり、図において、41は、各処理実行装置12の処理能力値を参照して、現在処理負荷が最小の処理実行装置12を検索するマネージャ、42は、各処理実行装置12に対応して備えられ、担当する処理実行装置12の処理能力値を参照し、タスクの処理を実行するために必要な機能が現在実行できる状態にあるか否かを判定するノードである。

【0026】図3は各ノード42の構成を示すブロック図であり、図において、51は各種メッセージの送受信を実施する通信処理部、52は担当する処理実行装置12の処理能力値を格納する処理能力データベース、53は処理能力データベースに格納されている担当する処理実行装置12の処理能力値に基づき、その処理実行装置12での要求されたタスク処理の可能性を判断する状況分析部である。

【0027】次に動作について説明する。図4は処理実行装置12の処理能力値算出部36が算出する処理能力値の更新処理を示すフローチャートである。処理能力値は処理実行装置12の処理状況を示すデータであり、現在処理中のタスク名、その処理に用いている機能名、処

理に用いている装置内のメモリの占有率と演算能力の消費率等の計算機資源から構成されるデータである。

【0028】図4のステップST1において、処理実行装置12の稼働状況監視部34は、常時、処理実行部33の処理状態を監視し、ステップST2において、監視した稼働データを状態記録部35に記録する。ステップST3において、状態記録部35は、稼働状況監視部34から送られた稼働データと、その直前に記録された稼働データを比較し、状態変化を検知したときに、その状態変化を処理能力値算出部36に報告する。状態の変化がない場合には、ステップST1に戻って、稼働状況監視部34による処理状態の監視を続行する。

【0029】ステップST4において、状態変化の報告を受けた処理能力値算出部36は、状態記録部35に記録された稼働データより、処理中のタスク名、実行中の機能名と関連付けた最新の処理能力値を算出する。ステップST5において、処理制御部32が、ネットワークを介して、管理装置11における担当のノード42に最新の処理能力値を送信する。

【0030】図3に示す各ノード42は、送信された最新の処理能力値を、通信処理部51を介して処理能力データベース52に格納する。このようにして、各ノード42は、常時、担当する処理実行装置12の最新の処理能力値を把握している。

【0031】図5は分散処理システムの全体の処理の流れを示すフローチャートである。ステップST11において、管理装置11のタスク受付部21は、外部の入力装置13よりタスク処理の命令を受け付け、ステップST12において、タスク受付部21は、受け付けたタスクを直ちに協調処理部23に出力すると同時に、そのタスク情報をタスクデータベース22に格納する。

【0032】ステップST13において、協調処理部23は、タスク受付部21からタスクの処理命令を受けると、マネージャ41が、各ノード42の状況分析部53に対して、タスクに対応した一意的オーナーシップの譲渡要求を行う。ここで、オーナーシップとは、タスクの処理実行管理権のことで、オーナーシップを所有することにより、タスクの処理を実行することが可能になる権利である。何れかのノード42がこのオーナーシップを所有することにより、そのノード42が担当している処理実行装置12にタスク処理の指示が可能になる。

【0033】オーナーシップの情報は、タスク名称と保有者名(ここでは、ノード名もしくは処理実行装置名)、そして処理状況から構成される。さらに、オーナーシップには、処理結果等のデータを添付することを可能とし、各ノード42間の処理結果を含めた情報伝達を実現する。

【0034】ステップST14において、各ノード42の状況分析部53が、通信処理部51を介してマネージャ41からのオーナーシップの譲渡要求を入手すると、

10

20

30

40

50

状況分析部53は、処理能力データベース52に格納されている最新の処理能力値に基づき、タスク処理の可否判断を行い、その判断結果と最新の処理能力値をマネージャ41に回答する。そして、マネージャ41は、各ノード42からの回答に基づき、オーナーシップの譲渡先を決定し、その決定結果を各ノード42に通知する。すなわち、このオーナーシップが割り当てられているタスクの処理を行う処理実行装置12が決定される。

【0035】ステップST15において、オーナーシップを獲得したノード42は、オーナーシップの情報を検索キーとして、タスクデータベース22からタスク情報を検索する。ステップST16において、ノード42は、検索したタスク情報からそのタスクの処理依頼を行うための通信メッセージを作成し、通信処理部51、通信処理部26を介して担当する処理実行装置12に送信する。

【0036】ステップST17において、処理実行装置12の処理制御部32は、管理装置11のノード42から通信メッセージを受け取り、タスクの処理指示と処理内容を分析し、その結果を処理実行部33に送信し処理の実行を開始する。ステップST18において、処理実行の開始に伴い、図4に示す処理能力値の更新処理と同様に、稼動状況監視部34は処理実行装置12の稼動状況の変化を検知し、処理能力値算出部36が最新の処理能力値を算出し、ステップST19において、処理制御部32は、処理能力値算出部36が算出した最新の処理能力値を担当のノード42に送信する。この処理により、ノード42は担当の処理実行装置12の処理能力値の変化を逐次、把握することが可能となる。

【0037】ステップST20において、処理実行部33が依頼された処理を完了すると、処理制御部32は完了報告と処理結果を担当のノード42に送信し、処理能力値算出部36は処理完了に伴う処理能力値を更新し、処理制御部32は担当のノード42に送信する。そして、処理完了報告を受けたノード42は、マネージャ41に、完了報告を行うと共に処理結果を通知し、マネージャ41は、タスク受付部21を介して、処理結果をタスクの要求元である入力装置13に送信する。

【0038】ステップST21において、オーナーシップ管理部24は、ステップST20における完了処理実施の終了を確認した後、タスク処理の終了とオーナーシップ割り当ての終了に伴い、オーナーシップデータベース25に格納されているオーナーシップ情報を更新する。

【0039】図6は図5のステップST14におけるオーナーシップ譲渡先決定処理の流れを示すフローチャートである。ステップST31において、各ノード42の状況分析部53は、マネージャ41から発行されたオーナーシップの譲渡要求を入手し、ステップST32において、状況分析部53は、処理能力データベース52に

格納されている担当の処理実行装置12の最新の処理能力値に基づき、タスク処理の可否判断を行う。

【0040】状況分析部53は、この場合の判断基準として、各処理実行装置12に固有な処理実行可能閾値を設定し、処理能力データベース52に格納されている処理能力値から現在の計算機資源の占有率を求め、要求されたタスクを実行したときの予測占有率を計算し、現在の計算機資源の占有率と予測占有率との合計値が処理実行可能閾値を越えなければ、そのタスクの処理は可能と判断する。また、両占有率の合計値が処理実行可能閾値を越えた場合には、処理は不可能と判断する。

【0041】処理可能と判断した場合、状況分析部53は、マネージャ41に、「オーナーシップ譲渡可能」と、そのときの処理能力値を送信する。処理不可能と判断したときは、「オーナーシップ譲渡不可能」を送信する。

【0042】マネージャ41は、ステップST33において、全ノード42からの回答を集約し、ステップST34において、各ノード42からの処理能力値を参照し、タスクの処理に必要な機能を有している最も負荷の少ないノード42を選択し、オーナーシップの譲渡先として決定する。このとき、オーナーシップ管理部24はオーナーシップの譲渡情報をオーナーシップデータベース25に記録する。ステップST35において、マネージャ41は、全てのノード42にオーナーシップの譲渡先の決定結果を通知する。

【0043】また、ステップST34において、マネージャ41が、一度目の各ノード42からの報告では、オーナーシップの譲渡先を決定できない場合には、ステップST32に戻って、交渉プロトコルに従い、各ノード42に対して再判断を依頼して、再評価を繰り返しながら譲渡先を決定する。ステップST34において、要求されたタスク処理が膨大で、ネットワークに接続されている各処理実行装置12で処理が不可能な場合には、ステップST36において、タスク処理の要求元である入力装置13に対して処理を拒否する。

【0044】以上のように、この実施の形態1によれば、管理装置11の協調処理部23が各処理実行装置12の最新の処理能力値を把握しておき、タスクの処理要求があった場合に、管理装置11が、把握している各処理実行装置12の最新の処理能力値に基づき、負荷の最も少ない処理実行装置12を決定するので、ネットワークに接続されている全ての処理実行装置12にタスクの処理の可否判断を求める必要がなく、通信料を減らすことができるという効果が得られる。

【0045】実施の形態2. この実施の形態2における分散処理システムの構成及び協調処理部の構成は、それぞれ実施の形態1における図1、図2に示す構成と同等である。図7はこの発明の実施の形態2によるノード42の構成を示すブロック図であり、図において、54は

10

20

30

40

50

他の処理実行装置12の処理能力値により実行モデルを作成管理するモデリング部であり、その他の構成は実施の形態1の図3に示す構成と同等である。

【0046】この実施の形態は、オーナーシップを譲渡された処理実行装置12が、要求されたタスク処理を開始した後に、処理状況の変化に対応して、処理効率向上のためにタスク処理を行う処理実行装置12を変更するものである。

【0047】次に動作について説明する。図8はタスク処理を行う処理実行装置12の変更処理の流れを示すフローチャートである。ステップST41において、オーナーシップを譲渡された処理実行装置12が、要求されたタスク処理を開始する。

【0048】ステップST42において、マネージャ41は各ノード42に処理開始のアナウンスを行うと、オーナーシップを譲渡されなかったノード42は、オーナーシップを譲渡されたノード42からの処理能力値により、オーナーシップを譲渡され、タスクの処理を行っている処理実行装置12の処理状況を観測する。

【0049】処理状況を観測した結果、オーナーシップを譲渡されなかったノード42の処理実行装置12が、代わりにタスクの処理を行った方が効率が良いと判断した場合には、ステップST43において、マネージャ41にオーナーシップの変更要求を行い、マネージャ41は、現在のタスクの処理を行っている処理実行装置12のノード42に対して、オーナーシップの変更指示を行う。

【0050】ステップST44において、オーナーシップの変更指示を受けたノード42は、担当の処理実行装置12に対して、現在行っているタスク処理の中断を指示し、指示を受けた処理実行装置12の処理実行部33が現在の処理を中断し、処理結果を処理結果集積部37に送信して処理を停止する。ステップST45において、処理結果を受取った処理結果集積部37は、処理内容と処理結果を関連付けて中間データをまとめ、処理制御部32が通信処理部31を介して委譲先の処理実行装置12に送信する。

【0051】ステップST46において、中間データを受取った委譲先の処理実行装置12は、中間データの内容を確認し、実行可能のときは担当のノード42に変更処理の成功を報告する。ステップST47において、成功の報告を受けたノード42はマネージャ41に報告し、マネージャ41はオーナーシップ管理部24にオーナーシップの更新依頼を行い、ステップST48において、オーナーシップ管理部24は、オーナーシップデータベース25に格納されているオーナーシップデータの更新を行う。

【0052】ステップST49において、委譲元の処理実行装置12の処理結果集積部37は、保有している中間データを破棄し、ステップST50において、委譲先

の処理実行装置12は、送信された中間データに基づき、委譲されたタスクの処理を継続する。

【0053】上記ステップST46において、委譲先の処理実行装置12が中間データの内容を確認し、実行不可能のときは、ステップST51において、委譲先の処理実行装置12は変更失敗を担当のノード42に報告し、担当のノード42はマネージャ41にオーナーシップの変更不可を通知する。マネージャ41は、オーナーシップの変更指示を取り消し、今までオーナーシップを譲渡されていたノード42に対して、今までの処理を続行するよう指示する。

【0054】図9は図8のステップST42の処理状況観測処理の流れを示すフローチャートである。ステップST61において、最初にオーナーシップを獲得した処理実行装置12がタスク処理を開始したときに、マネージャ41は各ノード42に処理開始のアナウンスを行う。ステップST62において、処理開始のアナウンスを受けた各ノード42のモデリング部54は、アナウンスの内容分析を行ってタスク名を抽出し、さらに処理に必要な機能名をタスクデータベース22から検索する。

【0055】ステップST63において、各ノード42のモデリング部54は、検索した機能名より、自分が保有している機能が使われているとき、言い換えれば、担当している処理実行装置12がタスクを実行可能となるときに、処理を開始した処理実行装置12の処理能力値により実行モデルを作成し、ステップST64において、モデリング部54は、処理を開始している処理実行装置12のノード42から定期的に処理能力値を収集して、作成した実行モデルのモデル情報を更新する。

【0056】ステップST65において、実行モデルを作成、更新しているノード42は、処理を開始した処理実行装置12の処理能力値によりモデル情報を更新する度に、自己の処理能力値と実行モデルのモデル情報とを比較して処理の効率の評価を行う。評価は演算能力等の固有の処理能力、リソースの消費率から、処理効率、処理時間を比較し、タスクの内容、特性により評価項目を選択し、比較結果を算出する。このとき、現状の処理が適しているときは正常と評価し、ステップST64で、そのままモデル情報を更新する。

【0057】また、処理を開始した処理実行装置12の処理効率が低下し、評価しているノード42が担当している処理実行装置12に処理を変更するほうが良いと判断したときは、ステップST66において、評価しているノード42はオーナーシップの変更要求を行う。

【0058】以上のように、この実施の形態2によれば、オーナーシップが譲渡されていないノード42が、タスクを処理している処理実行装置12の処理状況を観測し、担当する処理実行装置12が代わってタスクの処理を行った方が良いと判断した場合には、オーナーシップの変更処理を行うことにより、処理実行装置12の動

的な負荷の変化に対応することができるという効果が得られる。

【0059】実施の形態3. この実施の形態3における分散処理システムの構成及び協調処理部の構成は、それぞれ実施の形態1における図1, 図2に示す構成と同等である。図10はこの発明の実施の形態3によるノード42の構成を示すブロック図であり、55はモデリング部54が処理を開始している処理実行装置12の処理能力値より作成した実行モデルのモデル情報を予測する予測部であり、その他の構成は、実施の形態2の図7に示す構成と同等である。

【0060】次に動作について説明する。各ノード42のモデリング部54は、実施の形態2と同様に、作成した実行モデルの状態の変化を分析する。予測部55は、分析結果より変化率、変化則を算出し、所定の時間内での処理能力値を算出する。予測部55は同様に所定時間内の自己の処理能力値の予測も行い、モデリング部54は所定の時間内における同時刻での処理能力値の比較を行うことにより、処理効率の評価を行う。そして、実施の形態2と同様に、処理効率向上のためにオーナーシップの変更、処理実行装置12の変更を行う。

【0061】以上のように、この実施の形態3によれば、オーナーシップが譲渡されていないノード42が、タスクを処理している処理実行装置12の処理状況を観測し、担当する処理実行装置12が代わってタスクの処理を行った方が良いと判断した場合には、オーナーシップの変更処理を行うことにより、処理実行装置12の動的な負荷の変化に対応することができるという効果が得られる。

【0062】実施の形態4. この実施の形態4における分散処理システムの構成及び協調処理部の構成は、それぞれ実施の形態1における図1, 図2に示す構成と同等である。図11はこの発明の実施の形態4によるノード42の構成を示すブロック図であり、56はシステム全体の負荷の判定を行う負荷判定部、57は負荷の偏りを調整する代替案を作成する負荷調整部であり、その他の構成は、実施の形態2の図7に示す構成と同等である。

【0063】次に動作について説明する。図12は実施の形態2における図8のステップST42の処理状況観測処理の流れを示すフローチャートである。ステップST71において、最初にオーナーシップを獲得した処理実行装置12がタスク処理を開始したときに、マネージャ41は各ノード42に処理開始のアナウンスを行う。

【0064】マネージャ41からアナウンスを受けた各ノード42のモデリング部54は、ステップST72において、アナウンスの内容を分析し、ステップST73において、システムを構成する処理実行装置12全ての各実行モデルを作成し、ステップST74において、各実行モデルのモデル情報を更新する。

【0065】ステップST75において、各ノード42

のモデリング部54が各実行モデルのモデル情報を更新する度に、負荷判定部56は、各実行モデルについて処理能力値を分析し、高負荷な処理実行装置12を調査する。負荷判定部56が、高負荷な処理実行装置12が存在し、システム内の負荷に偏りがあると判断したとき、ステップST76において、負荷調整部57は、高負荷の処理実行装置12に代り、タスク処理を行うことのできる代替の処理実行装置12を検索する。このとき、負荷調整部57は、処理機能と処理実行装置12の処理状況をもとに、負荷が低く、処理能力が高い処理実行装置12を選択して、負荷の偏りを調整する代替案を作成する。

【0066】ステップST77において、負荷判定部56は、この代替案を適用した結果を評価し、システム内の負荷状況が改善されると判断したときに、ステップST78において、ノード42はその代替案をマネージャ41に通知し、オーナーシップ変更を要求する。この代替案を適用しても改善されないと判断したときは、ノード42は代替案を破棄し、ステップST74のモデル情報の更新を継続する。オーナーシップ変更の要求後の処理は、実施の形態2と同様である。

【0067】以上のように、この実施の形態4によれば、オーナーシップが譲渡されていないノード42が、タスクを処理している処理実行装置12の処理状況を観測し、担当する処理実行装置12が代わってタスクの処理を行った方が良いと判断した場合には、オーナーシップの変更処理を行うことにより、処理実行装置12の動的な負荷の変化に対応することができるという効果が得られる。

【0068】実施の形態5. この実施の形態5における分散処理システムの構成及び協調処理部の構成は、それぞれ実施の形態1における図1, 図2に示す構成と同等である。図13はこの発明の実施の形態5によるノード42の構成を示すブロック図であり、58は現在自分が処理しているタスクを他の処理実行装置12に依頼し、処理管理を行う外部依頼部であり、実施の形態4の図11における負荷調整部57を、外部依頼部58に置き換えたものである。

【0069】次に動作について説明する。図14はこの発明の実施の形態5における各ノード42の処理の流れを示すフローチャートである。ステップST81において、マネージャ41がオーナーシップの譲渡を行い、ステップST82において、オーナーシップが譲渡された処理実行装置12がタスク処理を開始する。

【0070】ステップST83において、オーナーシップが譲渡されたノード42のモデリング部54は、他の処理実行装置12の処理能力値から実行モデルを作成する。また、ステップST84において、モデリング部54は、自己の処理能力値から自己の実行モデルを作成する。ステップST85において、モデリング部54

は、処理状況を監視して、他の処理実行装置12の実行モデル、自己の実行モデルのモデル情報の更新を行う。

【0071】ステップST86において、負荷判定部56はモデル情報更新時に負荷の状態を評価し、正常の場合には、ステップST85のモデル情報の更新を継続する。モデリング部54が作成した実行モデルの何れかが、効率低下、処理継続不可能等で、負荷判定部56が異常と評価した場合には、ステップST87において、外部依頼部58は、他の実行モデルを参照し、最も処理能力が高い他の処理実行装置12を検索する。このとき、管理装置11で管理するオーナーシップ情報の更新、変更はせずに、依頼を行うノード42と処理実行装置12の判断、管理において、処理の権限のみを委譲する。

【0072】ステップST88において、外部依頼部58は、検索結果をもとに処理の依頼を担当のノード42に行う。依頼を受けたノード42は処理能力値に基づき処理可否の判断を行い結果を回答する。処理依頼が失敗したときは、ステップST85のモデル情報の更新を継続する。処理依頼が成功したときは、ステップST89において、実施の形態2と同様にして、依頼元の処理実行装置12により処理済の中間データが、依頼先の処理実行装置12に送信され、依頼先の処理実行装置12が処理を開始する。

【0073】ステップST90において、依頼先の処理実行装置12が処理を完了すると、ステップST91において、処理結果集積部37が処理結果を集積し、ステップST92において、依頼先の処理実行装置12は処理結果を依頼元の処理実行装置12に送信する。ステップST93において、処理結果を受信した依頼元の処理実行装置12は、処理完了を担当のノード42に通知して完了処理を行い、割り当てられたタスクの処理を完了する。

【0074】以上のように、この実施の形態5によれば、オーナーシップが譲渡されたノード42が、他の処理実行装置12の処理状況を観測し、他の処理実行装置12が代わってタスクの処理を行った方が良いと判断した場合には、オーナーシップの変更要求を行わずに、処理実行装置12間で、タスク処理の依頼を自由に行うことにより、局所的な処理状況の変化に柔軟に対応することができるという効果が得られる。

【0075】

【発明の効果】以上のように、この発明によれば、管理装置は、逐次把握している各処理実行装置のタスクと関連付けた処理状況を示す処理能力値に基づき、受け付けたタスクの処理を実行する処理実行装置を決定して、タスクの実行管理権であるオーナーシップを設定し、決定した処理実行装置に、受け付けたタスクの処理依頼を行う協調処理部を備え、処理実行装置は、管理装置から依頼されたタスクの処理を実行する処理実行部と、処理実

行部によるタスクの処理における稼働データを計測する稼働状況監視部と、稼働状況監視部が計測した稼働データを記録する状態記録部と、状態記録部に記憶されている稼働データを参照し、所定の間隔で処理中のタスクと関連付けた最新の処理状況を示す処理能力値を算出する処理能力値算出部と、処理能力値算出部が算出した最新の処理能力値を管理装置に送信する処理制御部とを備えたことにより、ネットワークに接続されている全ての処理実行装置にタスクの処理の可否判断を求める必要がなく、通信料を減らすことができるという効果がある。

【0076】この発明によれば、オーナーシップを譲渡されなかったノードが、オーナーシップを譲渡されたノードからの処理能力値に基づき、オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理状況を観測し、オーナーシップを譲渡されなかった処理実行装置が代わりにタスクの処理を実行した方が処理効率が良いと判断した場合に、マネージャにオーナーシップの変更要求を行うことにより、処理実行装置の動的な負荷の変化に対応することができるという効果がある。

【0077】この発明によれば、オーナーシップを譲渡されなかったノードが、オーナーシップを譲渡されタスクの処理を実行している処理実行装置の処理能力値により実行モデルを作成し、作成した実行モデルのモデル情報を更新し、自己の処理能力値と比較して、処理効率の評価を行うモデリング部を備えたことにより、処理実行装置の動的な負荷の変化に対応することができるという効果がある。

【0078】この発明によれば、オーナーシップを譲渡されなかったノードが、モデリング部が作成した実行モデルのモデル情報の変化率、変化則を算出し、所定の時間内の処理能力値を予測すると共に、所定の時間内の自己の処理能力値の予測を行う予測部を備え、モデリング部が、予測部が予測した処理能力値を比較することにより、処理効率の評価を行うことにより、処理実行装置の動的な負荷の変化に対応することができるという効果がある。

【0079】この発明によれば、ノードが、システム全体の処理実行装置の処理能力値により、各実行モデルを作成し、作成した各実行モデルのモデル情報を更新するモデリング部と、モデリング部が作成した実行モデルのモデル情報を更新する度に、全実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、負荷判定部の判定結果に基づき、負荷を調整する代替案を作成する負荷調整部とを備え、負荷判定部が、負荷調整部により作成された代替案を評価し、マネージャにオーナーシップの変更を要求することにより、処理実行装置の動的な負荷の変化に対応することができるという効果がある。

【0080】この発明によれば、ノードが、オーナーシップを譲渡されなかった処理実行装置の処理能力値により、他の実行モデルを作成すると共に、自己の処理能力

10

20

30

40

50

値から自己の実行モデルを作成し、他の実行モデルのモデル情報と、自己の実行モデルのモデル情報を更新するモデリング部と、モデリング部が作成した各実行モデルの処理能力値を分析し負荷の判定を行う負荷判定部と、負荷判定部の判定結果に基づき、負荷を調整するために処理能力の高い処理実行装置に対応するノードにタスク処理の依頼を行う外部依頼部とを備え、処理実行装置間でタスク処理の依頼を自由に行うことにより、局所的な処理状況の変化に柔軟に対応することができるという効果がある。

【図面の簡単な説明】

【図1】 この発明の各実施の形態による分散処理システムの構成を示すブロック図である。

【図2】 この発明の各実施の形態による協調処理部の構成を示すブロック図である。

【図3】 この発明の実施の形態1によるノードの構成を示すブロック図である。

【図4】 この発明の各実施の形態による処理能力値算出部が算出する処理能力値の更新処理を示すフローチャートである。

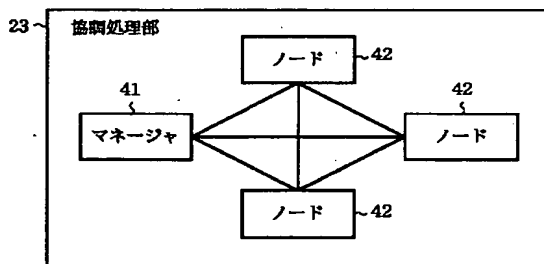
【図5】 この発明の実施の形態1による分散処理システムの全体の処理の流れを示すフローチャートである。

【図6】 この発明の実施の形態1によるオーナーシップ譲渡先決定処理の流れを示すフローチャートである。

【図7】 この発明の実施の形態2によるノードの構成を示すブロック図である。

【図8】 この発明の実施の形態2によるタスク処理を行う処理実行装置の変更処理の流れを示すフローチャートである。

【図2】



トである。

【図9】 この発明の実施の形態2による処理状況観測処理の流れを示すフローチャートである。

【図10】 この発明の実施の形態3によるノードの構成を示すブロック図である。

【図11】 この発明の実施の形態4によるノードの構成を示すブロック図である。

【図12】 この発明の実施の形態4による処理状況観測処理の流れを示すフローチャートである。

10. 【図13】 この発明の実施の形態5によるノードの構成を示すブロック図である。

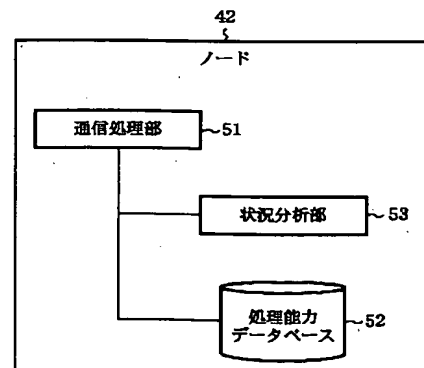
【図14】 この発明の実施の形態5によるノードの処理の流れを示すフローチャートである。

【図15】 従来の分散処理システムの構成を示すブロック図である。

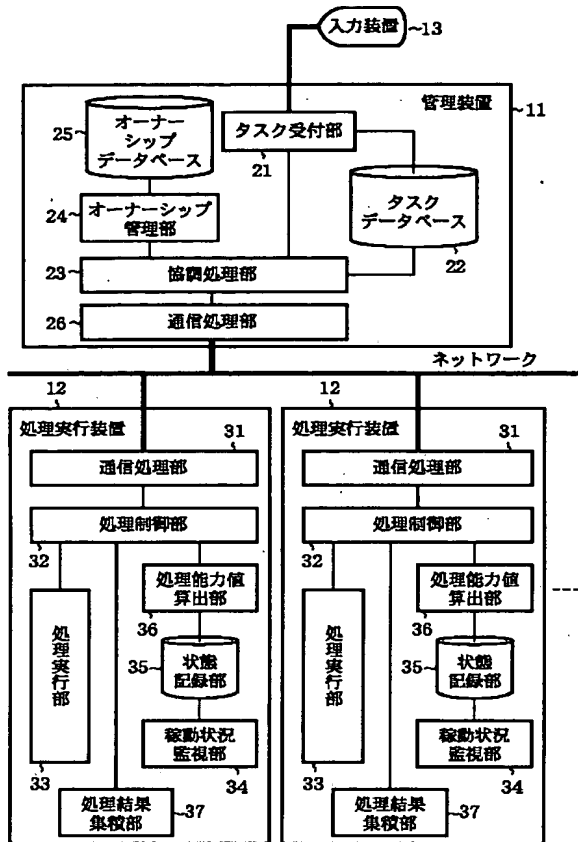
【符号の説明】

11 管理装置、12 処理実行装置、13 入力装置、21 タスク受付部、22 タスクデータベース、23 協調処理部、24 オーナーシップ管理部、25 オーナーシップデータベース、26 通信処理部、31 通信処理部、32 処理制御部、33 処理実行部、34 稼動状況監視部、35 状態記録部、36 処理能力値算出部、37 処理結果集積部、41 マネージャ、42 ノード、51 通信処理部、52 処理能力データベース、53 状況分析部、54 モデリング部、55 予測部、56 負荷判定部、57 負荷調整部、58 外部依頼部。

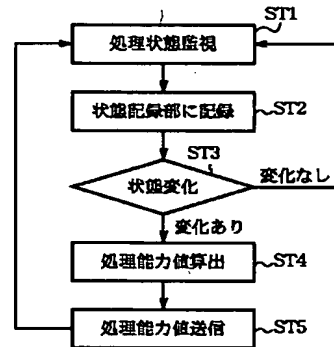
【図3】



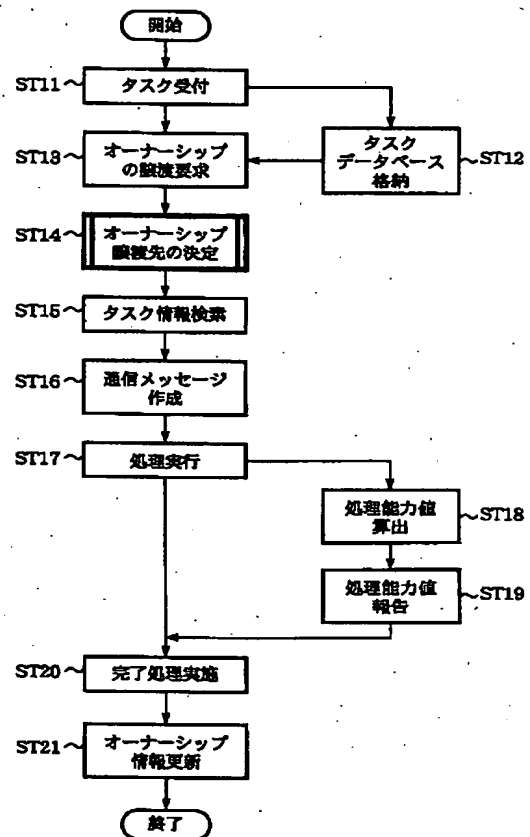
【 図1 】



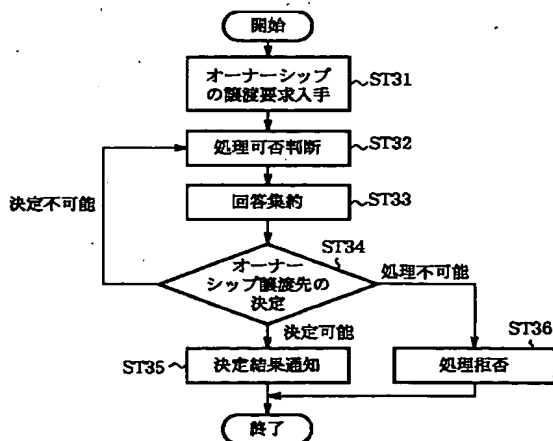
【 図4 】



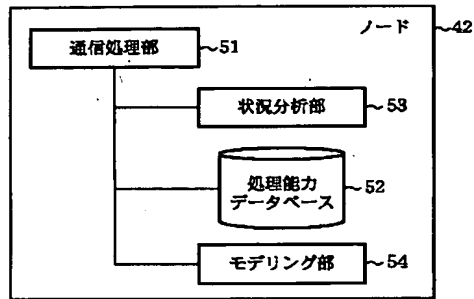
【 図5 】



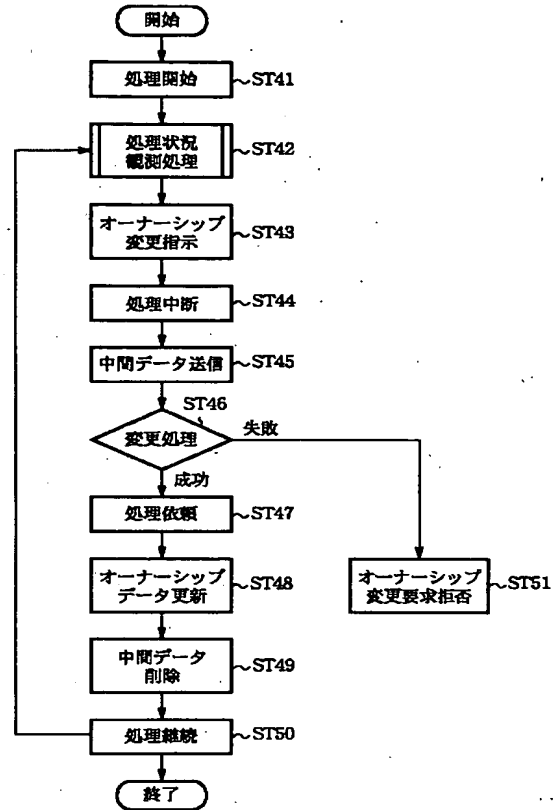
【 図6 】



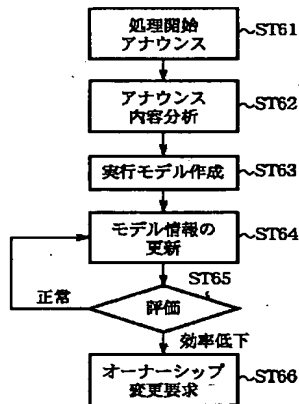
【 図7 】



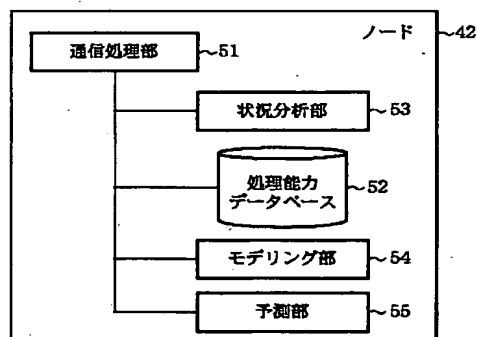
【 図8 】



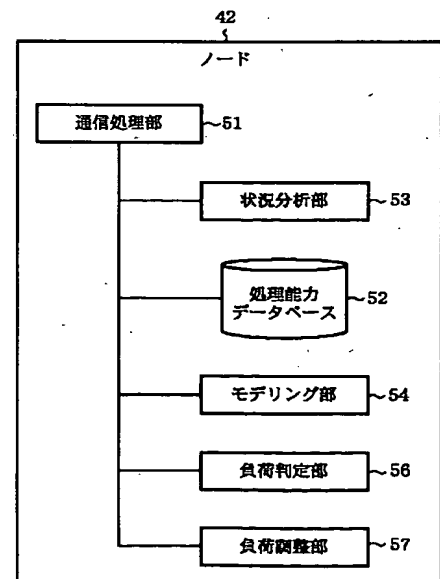
【 図9 】



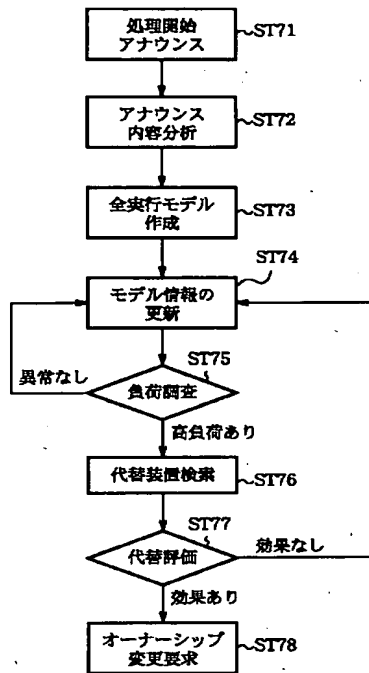
【 図10 】



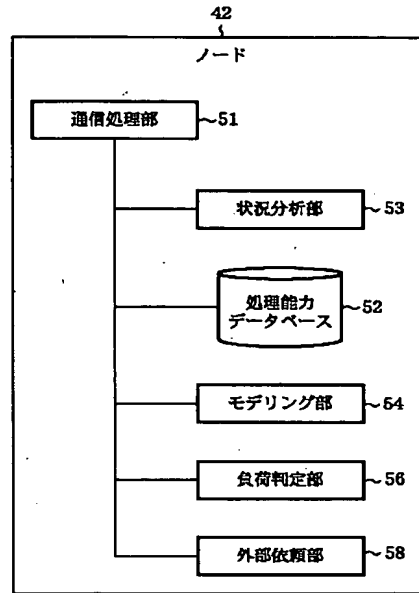
【 図11 】



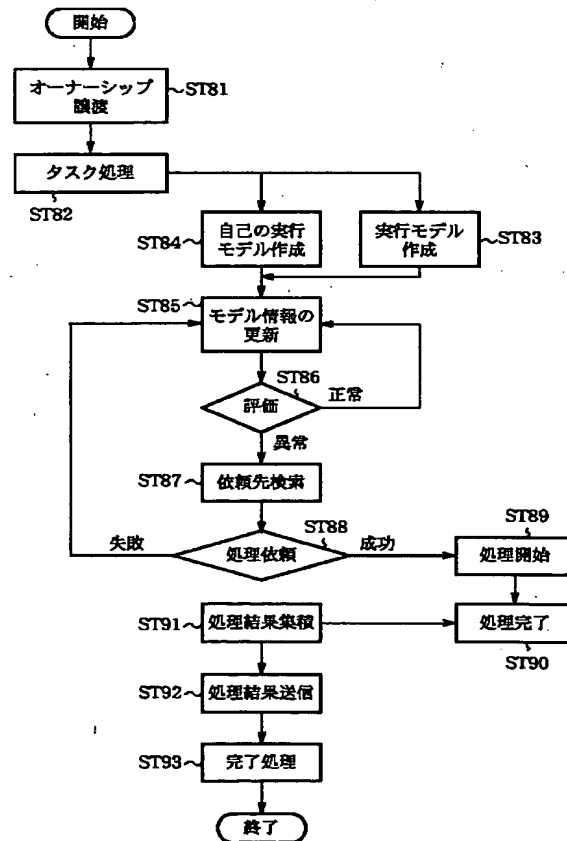
【 図12 】



【 図13 】



【 図14 】



【 図15 】

